

Our Ref. No.: 51876.P246
Express Mail No. EL651891033US

UTILITY APPLICATION FOR UNITED STATES PATENT

FOR

**INFORMATION GENERATION AND RETRIEVAL METHOD BASED ON
STANDARDIZED FORMAT OF SENTENCE STRUCTURE AND SEMANTIC
STRUCTURE AND SYSTEM USING THE SAME**

Inventor(s): Keon-Hoe Cha et al.

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
220

INFORMATION GENERATION AND RETRIEVAL METHOD BASED ON
STANDARDIZED FORMAT OF SENTENCE STRUCTURE AND
SEMANTIC STRUCTURE AND SYSTEM USING THE SAME

5 Field of the Invention

10 The present invention relates to an information
generation and retrieval apparatus based on a standardized
format of sentence structure and semantic structure and a
method thereof and a computer readable recording medium for
15 recording a program for implementing the method; and, more
particularly, to an information generation and retrieval
apparatus based on a standardized format of sentence structure
and semantic structure for describing and retrieving
20 information by generating natural language sentence that is
relevant to a standardized format for sentence structure and
semantic structure in order to generate, retrieve and
transport information and knowledge fast and efficiently on
the Internet and a method thereof and a computer readable
recording medium for recording a program for implementing the
method.

Prior Art of the Invention

25 One of most important factor is establishment of
information services capable of storing, managing and
transporting mass data in forwarding to information society.

Digitalized data stored in databases and networks allowing remote online accessing evoke an essential technical issue of information management and information retrieval. Information retrieval is partitioned into two parts, information processing part in which information is accumulated and processed and information presentation part in which information is presented to an information requester and searched. And also, there is included a user interface technique for helping the information requester to use the information services more conveniently and efficiently.

Relating to information brokerage, multi-national PC communication companies like AOL(American On-Line) in U.S.A. provide a variety of information transport services and there is a trend of cooperating with Web information retrieval engine companies (e.g., Yahoo) to widening their commercial categories, however, there has no technique for brokering information commerce candidate adequately and fast and meditating commerce based on semantic representation of sentence and semantic likelihood.

Relating to semantic representation technique, information retrieval and natural language processing research centers like CIIR(Center for Intelligent Information Retrieval) in Massachusetts University in U.S.A. has developed an applied technique employing exquisite information extraction and information retrieval techniques based on word concept and semantic analysis of information from 1998, however, there has been no effort to apply them to information

commerce. And also, Webpage markup language using OML(Ontology Markup Language), CKML(Conceptual Knowledge Markup Language) and so on is introduced in University of Maryland and. University of Southern California. Research for
5 "Mikro-Kosmos" in New Mexico state University, "WordNet" in Princeton University, "CYC" of Cycorp, KIF(Knowledge Interchange Format) and Ontology for sharing reuse are progressed and each of institutes makes effort to suggest its scheme as a standard and, recently, IEEE(Institute of
10 Electrical and Electronics Engineers) makes effort to establish a generalized and abstract upper level concept by standardized semantic hierarchical structure referred as SUO(standard Upper Ontology).

In Korea, PC communication companies such as Chollian,
15 Naunuri, Hitel and so on relating to information transport determine commerce capability of an information providing commerce applicant and give information provider license to a selected applicant to allow it to collect and process information monopolisticly and register adequate menu and
20 information PC communication community, and charge the bill for time for reading information to users. And also, Information Transaction Center set up in 1998 to handle information transaction through the Internet and already started brokering of Patent technique area. However, there is
25 no protocol of information transaction and no concept of likelihood-based information extraction/matching technique for selecting proper transaction candidate automatically

accurately is employed so that mass transaction of information

Relating to semantic representation technique, Cheonbuk University and Pohang University of Science and Technology in Korea study basics for conceptual graph and ETRI, Ulsan University, Hannam University, Incheon University and Hoseo University study thesaurus systems. However, standardization of semantic representation and related techniques has not been established yet.

Accordingly, in the art of studying a standardized format for sentence structure and semantic structure, it is required a technique capable of describing and retrieving information by generating natural language sentence that is relevant to a standard for sentence structure and semantic structure in order to generate, retrieve and transport information and knowledge fast and efficiently on the Internet.

Summary of the Invention

Therefore, it is an object of the present invention to provide an information generation and retrieval apparatus based on a standardized format for sentence structure and semantic structure of natural language sentences for user's describing and retrieving information, a method thereof and a computer readable recording medium for recording a program for implementing the method.

In accordance with an aspect of the present invention, there is provided an apparatus for generating and retrieving

information based on standardized formats of sentence structure and semantic structure, the apparatus comprising;

a data storing unit for storing language knowledge data used to analyze a sentence for information supply and a query
5 for information request from a user, semantic representation data for representing sense of sentence as a conceptual graph, and Web documents;

an input unit for receiving a natural language query sentence for generation of a natural language sentence for
10 information supply and specification of information request from the user;

an input sentence analyzing unit for analyzing sentence structure of the natural language sentence or the natural language query sentence inputted from the user with reference
15 to data stored at the data storing unit to generate semantic structure;

semantic structure processing unit for partitioning the semantic structure analyzed by the input sentence analyzing unit to index and store or for computing semantic relevance to
20 search supply information and document most semantically relevant to the requested information specification;

an interactive processing unit for outputting sentence format rule for which failure data from the input sentence analyzing unit is corrected depending on the standardized
25 formats of sentence structure and semantic structure, and indexing and searching result; and

an information transferring unit for transferring the

data from the interactive processing unit to the user.

In accordance with another aspect of the present invention, there is provided a method for generating and retrieving information for use in an apparatus for generating and retrieving information based on standardized formats of sentence structure and semantic structure, the method comprising the steps of:

(a) transforming a natural language sentence (information and knowledge) described by a information provider to a conceptual graph depending on standardized formats of sentence structure and semantic structure and indexing the conceptual graph; and

(b) transforming a natural language query sentence inputted from a user to a conceptual graph depending on the standardized formats of sentence structure and semantic structure and searching information relevant to the requirement of the user among the indexed information.

In accordance with still another aspect of the present invention, there is provided an information generating method for use in an information generating apparatus based on standardized formats of sentence structure and semantic structure, the method comprising the steps of:

(a) generating a sentence in which ambiguities in sentence structure and semantic structure are solved depending on the standardized formats of sentence structure and semantic structure from a natural language sentence inputted by a information provider;

(b) transforming the generated sentence to a conceptual graph by sentence analysis and semantic analysis; and

(c) transforming the transformed conceptual graph to a record of a table by a relation node and indexing the record.

5 In accordance with still another aspect of the present invention, there is provided an information retrieving method for use in an information retrieving apparatus based on standardized formats of sentence structure and semantic structure, the method comprising the steps of:

10 (a) analyzing sentence structure and semantic structure of a natural language query sentence received from a user to transform it to a conceptual graph;

(b) searching a conceptual graph in a database semantically nearest to the conceptual graph of the query and
15 computing semantic relevance; and

(c) retrieving indexed information by the searched conceptual graph and provide it to the user.

In accordance with still another aspect of the present invention, there is provided a computer readable medium for
20 recording a program for implementing, at an information generating and retrieving apparatus based on standardized formats of sentence structure and semantic structure having a processor, the functions of:

(a) transforming a natural language sentence (information
25 and knowledge) described by a information provider to a conceptual graph depending on standardized formats of sentence structure and semantic structure and indexing the conceptual

graph; and

(b) transforming a natural language query sentence inputted from a user to a conceptual graph depending on the standardized formats of sentence structure and semantic structure and searching information relevant to the requirement of the user among the indexed information.

In accordance with still another aspect of the present invention, there is provided a computer readable medium for recording a program for implementing, at an information generating apparatus based on standardized formats of sentence structure and semantic structure having a processor, the functions of:

(a) generating a sentence in which ambiguities in sentence structure and semantic structure are solved depending on the standardized formats of sentence structure and semantic structure from a natural language sentence inputted by a information provider;

(b) transforming the generated sentence to a conceptual graph by sentence analysis and semantic analysis; and

(c) transforming the transformed conceptual graph to a record of a table by a relation node and indexing the record.

In accordance with still another aspect of the present invention, there is provided a computer readable medium for recording a program for implementing, at an information retrieving apparatus based on standardized formats of sentence structure and semantic structure having a processor, the functions of:

(a) analyzing sentence structure and semantic structure of a natural language query sentence received from a user to transform it to a conceptual graph;

(b) searching a conceptual graph in a database semantically nearest to the conceptual graph of the query and computing semantic relevance; and

(c) retrieving indexed information by the searched conceptual graph and provide it to the user.

Brief Description of the Drawings

The above and other objects and features of the instant invention will become apparent from the following description of preferred embodiments taken in conjunction with the accompanying drawings, in which:

Fig. 1 is a block diagram of an information generation and retrieval apparatus based on a standardized format of sentence structure and semantic structure in accordance with the present invention;

Figs. 2A and 2B show flow charts for an embodiment of the information generation and retrieval method based on a standardized format of sentence structure and semantic structure in accordance with the present invention;

Fig. 3 is a detailed flow chart for the embodiment of the information generation and retrieval method as shown in Figs. 2A and 2B;

Figs. 4A and 4B are exemplary diagrams of sub-graph

generated by sentence category and conversion rule and concept processing result depending on the sentence category in accordance with the present invention;

Figs. 5A to 5F are exemplary diagram of table structure and record contents generated by division graph indexing result in accordance with the present invention; and

Fig. 6 is a flow chart for an embodiment of user request information and document retrieval procedure in accordance with the present invention.

Preferred Embodiment of the Invention

The present invention comprises information supply procedure in which information and knowledge described by a user are transformed to a concept graph to be stored and indexed by generating natural language sentences depending on standardized formats for sentence structure and semantic structure and information request in which information adequate to a user's requirement by analyzing user's query is retrieved and extracted.

During the information supply procedure, a sentence for which ambiguities of sentence structure and semantic structure is solved by a toolset for supporting information specification for facilitating generation of information and knowledge depending on standardized formats of sentence structures and semantic structure and then the generated sentence is transformed to a conceptual graph by analyzing the

sentence structure and the semantic structure. And the conceptual graph is transformed to a record value in a table by relation node, to be stored and indexed.

During the information request procedure, user's natural language query is transformed to the conceptual graph by analyzing sentence structure and semantic structure and then a conceptual graph in a database, which is nearest to the conceptual graph of the query with respect to sense is searched and semantic appropriateness is computed to display information indexed by the searched conceptual graph to the user.

That is, in the present invention, information and knowledge can be retrieved and transacted efficiently and accurately by solving ambiguities in sentence structure and semantic structure by interacting with the user from the step of generating information and knowledge of the user for efficient information transaction by a standard format of sentence structure and semantic structure and by matching specification information of the user to the standardized format of the sentence structure and the semantic structure.

Hereinafter, preferred embodiments of the present invention will be described in detail with reference to the accompanying drawings.

Fig. 1 is a block diagram of an information generation and retrieval apparatus based on a standardized format of sentence structures and semantic structure in accordance with the present invention.

As shown in Fig. 1, the information generation and retrieval apparatus based on the standardized format of sentence structure and semantic structure of the present invention comprises a data storage part 11 for storing language knowledge data to be used for analyzing a sentence for supplying information from a user and a query for requesting information, semantic representation data for representing sense of the sentence by a conceptual graph, and Web documents, an information supply interface 12 for receiving natural language query for generating the sentence for supplying information to be transacted or specifying information request from the user, an input sentence analyzing part 13 for generating semantic structure by the inputted natural language sentence from the user with referring to the data at the data storage part 11, a interactive response processor 14 for solving spelling error or spacing error happened in the sentence generated by the user from the input sentence analyzing part 13, non-sentence disobeying the standardized formats for sentence structures and semantic structure and ambiguities in sentence structure and semantic structure to output the solved result and indexing and retrieving results, a semantic structure processor 15 receiving the semantic structure from the input sentence analyzing part 13 form indexing/storing or retrieving, and an information demand interface 16 for providing data from the interactive response processor 14 to the user.

Herein, the data storage part 11 includes a lexicon

storage unit 111, a predicate case frame storage unit 112, a noun thesaurus storage unit 113, a concept graph database 114, and a Web document database 115. The input sentence analyzing part 13 includes a morphological analyzer 131, a parser 132 and a semantic structure generator 133. The interactive response processor 14 includes an analysis failure processor 141 and a response generator 142. The semantic structure processor 15 includes a conceptual graph transformer 151, a conceptual graph indexer 152 and a conceptual graph searcher 153.

It will be described in detail for operation of the information generation and retrieval apparatus based on the standardized format of sentence structure and semantic structure of the present invention as described above.

Firstly, when the information supply interface 12 receives information by the natural language sentence or the information demand interface 16 receiving specification for information required from the user transfers the inputted sentence to the input sentence analyzing part 13, the input sentence analyzing part 13 sequentially analyzes the inputted sentence by comparing it with data at the lexicon storage unit 111, the predicate case frame storage unit 112, and a noun thesaurus storage unit 113 of the data storage part 11. At that time, the morphological analyzer 131 does morphological analysis, the parser 132 parses the sentence to generate sentence structure tree, and the semantic structure generator 133 performs semantic analysis to generate the semantic

structure, i.e., conceptual graph.

The conceptual graph transformer 151 discriminates the conceptual graph from the semantic structure generator 133 depending on semantic relation and the conceptual graph
5 indexer 152 indexes Web documents including the supplied information from the user by using record of conceptual pair connected each discriminated relation. The conceptual graph searcher 153 searches the supplied information of highest semantic relevance by computing semantic relevance between
10 semantic structure of the user's query and the stored semantic structure.

The analysis failure processor 141 receives failure type data produced for analysis failure happened by disobedience of the user against standard rule during analysis of the
15 morphological analyzer 131, parser 132 and the semantic structure generator 133 and finds out corrected sentence format rule matched to the standard formats for sentence structure and semantic structure. The response generator 142 receives the corrected sentence format rule from the analysis
20 failure processor 141, retrieval result from the conceptual graph searcher 152 or indexing result from the conceptual graph indexer 153 and produces response type and contents for the user.

The standardized formats for sentence structure and
25 semantic structure can be expressed as follows.

(1) Standardized format for sentence structure

201. noun paragraph

NP1: (_NP+jc1)

NP2: (_NP+jc2)

NP3: (_NP+jc3)

202. Basic sentence

5 S: (NP1 _VP)

S: (NP1 NP2 _VP)

S: (NP1 NP3 _VP)

S: (NP1 NP2 NP3 _VP)

203. noun phase

10 NP: (nclnn+[nclnn]+[nclnn])

NP: (nblnp)

204. embellishing word extended noun paragraph

_NP: (mm! [dt]+[mm]!. [nulad] NP!-[nn,nb])

205. noun paragraph extension of conjunctive auxiliary

15 word, conjunctive adverb

_NP: (NP+jj NP)

_NP: (_NP maj _NP)

_NP: (_NP!@+', ' maj NP)

206. noun paragraph extension of the genitive case

20 _NP: (_NP+jm _NP!^[NP])

207. descriptive word

VP: (pv+[ep]+[ef])

VP: (pa+[ep]+[ef])

VP: (NP+co+[ep]+[ef])

25 VP: (NP+xsv+[ep]+[ef])

208. descriptive word extension of adverb

VP: (mag _VP!^[VP])

VP:(maj _S)

209. descriptive word extension of auxiliary declinable
word

VP:(_VP:(_VP+ec px+[ec] [px]+[ep]+[ef])

5 210. noun paragraph extension of embellishing clause

_NP:(ETMS:(_S+etm) NP)

211. noun clause

_NP:(_S+etn)

_NP:(_S+etm 'gut')

10 212. descriptive word extension of quotation clause

VP:(JQTS:(_S+'lago(go) hago'_VP))

VP:(JQTS:('''+_S+'''+'lago(go) hago'_VP))

213. descriptive word extension of adverb clause

VP:(ADVS:(NP1 [NP2] [NP3] 'ge|deutsi|dofok') _VP)

15 VP:(ADVS:(NP1 [NP2] [NP3] 'gati|eobsi|dali') _VP)

214. description clause

VP:(NP1 _VP!^[VP])

215. conjunction

S:(_S+ec _S)

20 (2) Standardized format for semantic structure

<conceptual relation>

216. agent

'yi/ga', 'eun/neun', '~ggeseo', '~eseo', '~yeu'

217. experiencer

25 'yi/ga', 'eun/neun', '~yeu'

218. object

'eul/reul', '~wa/gwa', 'eseo', '~yeu'

219. location

'~eul', '~e', '~eseo', '~(eu)ro'

219a. starting point

219b. target point

5 220. C1(TIME)C2

'~e'

220a. starting point

220b. target point

221. purpose

10 '~eul', 'ryumyun'

222. causal-effect

'~(eu)ro'

223. receiver

'~ege', '~e'

15 224. result

'~(eu)ro'

225. instrument

'~ro', '~e', '~a/eo(seo)'

226. comparison

20 '~wa/gwa', '~mankeum', '~boda', 'churum'

226a. comparison

226b. equivalence

226c. difference

226d analogy

25 227. property

'~yeu', '#' [# is space]

227a. owner

227b. source

227c. location

228d. belonging to

229e. substance

5 229f. reciprocal

229g. subsumption

229h. whole-part

29i. property of

229j. juxtapose

10 <conceptual graph relation>

230. Juxtaposition

231. condition

232. causal-effect

233. context switch

15 233a. unfinished precedence

227b. finished precedence

234. involvement

235. possibility

236. concession

20 Next, It will be described in detail for operation of
each of the elements with referring to Figs. 2a to 6.

Figs. 2A and 2B show flow charts for an embodiment of the
information generation and retrieval method based on
standardized formats of sentence structure and semantic
25 structure of the present invention.

As shown in Figs. 2A and 2B, in order to generate and
retrieve information based on the standardized formats of

sentence structure and semantic structure in accordance with the present invention, at step 201, the user inputs natural language sentence for providing information or required information through information supply interface 16. At step 5 202, the morphological analyzer 131 analyzes the inputted natural language sentence morphologically.

Subsequently, at step 203, it is checked whether morphological analysis is performed successfully and, at step 204, if morphological analysis fails, failure type data are generated depending on type of morphological analysis failure and, at step 205, if morphological analysis is performed successfully, parsing is performed by using the morphological analysis result. 10

Then, at step 206, it is checked whether sentence structure relevant to information transaction standard is generated and, at step 207, if generation of the sentence structure fails, parsing failure type data is generated and, at step 208, if the sentence structure is generated successfully, parse tree is transformed to generate the sentence structure. 15 20

Subsequently, at step 209, it is checked whether semantic structure relevant to the semantic structure standard is generated and, at step 210, if generation of the semantic structure fails, semantic structure generation failure type data are generated and, at step 211, if the semantic structure is generated successfully, the semantic structure is inputted to the conceptual graph transformer 151 to partition the 25

conceptual graph.

And then, at step 212, it is checked whether the sentence inputted by the user is for information supply and, at step 213, if it is for information supply, each of the partitioned
5 conceptual graph of the sentence is indexed for relation node. And, at step 214, data is generated depending on conceptual graph indexing result and, at step 215, response format and contents of an information transaction system are generated to output final response at step 216. On the other hand, at step
10 217, if the sentence inputted by the user is a query for information request, the conceptual graph nearest to the conceptual graph stored at the conceptual graph database 114 semantically is retrieved to generate conceptual graph retrieval result at step 218. And then, the procedure is
15 progressed to the step 215 of generating the response format and contents of the information transaction system.

On the other hand, after generating the morphological analysis failure type data at step 204, generating the parse failure type data at step 207 and generating the semantic
20 structure generation failure type data, the analysis failure processor 141 generates failure correction rule data at step 219 and then the procedure is progressed to the step 215 of generating the response format and contents of the information transaction system.

25 It will be described in detail for operation of an information generating and retrieving method based on the standardized formats of sentence structure and semantic

structure of the present invention as described above.

The information generating and retrieving method based on the standardized formats of sentence structure and semantic structure comprises the steps of receiving a sentence through
5 a text editor, morphologically analyzing the inputted sentence, presenting the analysis result to the user if morphological ambiguities or spelling error and receiving a corrected sentence from the user.

After receiving the morphological analysis result of the
10 sentence and parsing, the analysis result is presented to the user if morphological ambiguities or spelling error and then the corrected sentence is received from the user or a number of sentence structures are provided to the user to let the user input a sentence matched to the standardized format of
15 sentence structure.

And then, after receiving parsing result of the sentence and generating the semantic structure, if the semantic structure is not matched to the standardized format of the semantic structure or ambiguities of semantic structure
20 happens, a number of semantic structures are provided to the user to let the user select one of them in order to match the sentence inputted by the user to the standardized format of semantic structure. And, partitioning and storing the semantic structure generated by semantic analysis result of
25 the sentence with respect to the semantic relation node, if the user's information is for information supply, it is indexed by a record including the semantic relation node and

its correlated conceptual node pair and identification of the document to be stored.

If the user's information is a specification for information request, semantic relevance between the semantic structure of the user's query and the semantic structure for supply information stored is computed so as to search semantically nearest supply information. And, receiving processing result from the analysis failure processor and the semantic structure processor, format and contents of result to be presented to the user are determined and response to be presented to the user is generated to output the response.

That is, the information generating and retrieving apparatus for information transaction of the present invention makes the sentence inputted by the user matched to (1) the standardized format of sentence structure (2) the standardized format of semantic structure and finally indexes to the conceptual graph for information supply procedure and analyzes the user's query and computes the semantic relevance by using the partitioned conceptual graph to search the nearest information semantically for information request procedure. The procedures are described for an exemplary sentence "a method for buying a good man perfume at Paris".

(A) sentence structure failure type data

(0,306,310)

failure type 0: standard disobedience failure

. (failure type number, (standard

disobedience identifier)+)

failure type 1: sentence structure ambiguities

(failure type number, (sentence structure)+)

5

(B) user presentation sentence structure type

((joeun (namja (hyangsu))reul) (parieseo) (((sal) su) issneun) bangbub) (a method for buying a good man perfume at Paris)

10

((joeun namja) hyangsu)reul) (parieseo) (((sal) su) issneun) bangbub) (a method for buying a good man perfume at Paris) (X)

15

As described above, for the sentence "joeun namja hyangsureul parieseo saneun bangbub (a method for buying a good man perfume at Paris)", morphological analysis result is generated without failure after the morphological analysis procedure.

20

However, undergoing the sentence analysis procedure, sentence structure failure type data as (A) is generated through the step 207 by the analysis failure processor 141 for the sentence by the restriction rule of the noun paragraph extension of the genitive case and the noun paragraph extension of the embellishing clause and the predicate case and frame information.

25

The (A) presents an example of failure type data of the sentence and typical data format of the failure type data.

The failure type data is inputted to the response generator 142 and undergoes the step of response generating of the information transaction system and presents the result as the (B) to the user so that the user can select one among presented sentence structures to describe the sentence again.

The sentence structure tree of the input sentence that is selected by the user or is matched with the standardized sentence structure is transformed to the semantic structure by the semantic structure generator 133.

Fig. 3 is a detailed flow chart for the step of semantic structure generating of the information generation and retrieval method as shown in Figs. 2A and 2B.

As shown in Fig. 3, during the step of generating the semantic structure, a sentence tree (T) in which the sentence structure ambiguities are solved is inputted from the semantic structure generator 133 at step 301, and the sentence tree is transformed to a P-CG (pre-Conceptual Graph) in which the semantic ambiguities are not solved at step 302 by searching the nodes of the sentence tree in order of left-right-parent node and combining sub-graph generated from left-right-child nodes. A 0-level non-terminal having no left-right-child node is transformed to the sub-graph by using a tree transformation rule. The tree transformation rule is shown in Fig. 4A, which presents each sentence category used at the parser and its corresponding transformation rule.

For the P-CG 303 generated as described above, conceptual node processing step 303 and conceptual node relation

determining step S304 are processed in order to transform it to the conceptual graph in which semantic ambiguities is solved.

Herein, at the conceptual node processing step 303,
5 information to be processed as a referent is searched within the P-CG by using a definitive processing rule and designated as the referent, and information such as a proper noun and the tense is set as type information of the concept by using thesaurus information. The sub-graph generated after
10 conceptual processing result for a conceptual node "pari" of the exemplary sentence "joeun namja hyangsureul parieseo saneun bangbub (a method for buying a good man perfume at Paris)" is shown in Fig. 4B. Subsequently, after the conceptual node processing, relation between the conceptual
15 nodes is determined at the step S304 by the thesaurus and frame information. Relation between a predicate case that is a verb or an adjective and a conceptual node in subcategory of the predicate case is determined by the predicate frame and a case auxiliary word. And relation among nouns in a composite
20 noun phrase is analyzed by using a rule and statistics.

When the semantic ambiguities happen during the procedure as described above, statistical word sense disambiguation is performed or a number of alternatives are presented to the user. For example, for "pari(Paris)" and "sa(buying)" in the
25 exemplary sentence "joeun namja hyangsureul parieseo sal su issneun bangbub(a method for buying a good man perfume)" having semantic ambiguities of [city:Paris], [insect:fly] and

[sa:live], [sa:buy]], respectively, the user is questioned to select one of them. After the procedure as described above, the conceptual graph of the final semantic structure is generated at the step S305.

5

(C) sentence structure tree : joeun namja hyangsureul parieseo sal su issneun bangbub

((joeun ((namja) hyangsu))reul) ((pari)eseo) (((((sa)l) su) iss)neun) bangbub)

10

(D) semantic structure conceptual graph : joeun namja hyangsureul parieseo sal su issneun bangbub

[bangbub]->(ATTR)->[Assertion030]

Assertion030: [sa]->(AGENT)->[NULL]

15

->(OBJECT)->[perfume:X]-

>(ATTR) [Assertion031]

->(JUXTPOSE) [man]

->(LOCATE)->[Paris]

->(ATTR)->[inexpensive]

20

Assertion031: [joeun]->(EXPERIENCER)->[X]

(E) semantic structure failure type data

(1, pari, Paris/city, fly/insect, sa, sa/live, sa/buy)

25

semantic structure failure type 0 : standard semantic structure disobedience

(0, (standard semantic structure disobedience item)+)

semantic structure failure type 1: semantic
ambiguities occurrence

(1, semantic ambiguities occurrence node identifier,
(semantic ambiguities occurrence node identifier sense)++)

5

(F) Partition graph

(ATTR) [bangbub] [Assertion030]

(AGENT) [sa] [NULL]

(OBJECT) [sa] [perfume:X]

10

(LOCATE) [sa] [Paris]

(ATTR) [sa] [inexpensive]

(ATTR) [X] [Assertion03']

(JUXTAPOSE) [X] [namja]

(EXPERIENCE) [joeun] [X]

15

The (C) and (D) present the sentence tree and the
conceptual graph of the semantic structure for the sentence
"joeun namja hyangsureul parieseo sal su issneun bangbub",
which are inputs of the semantic structure generating step 208.

20

For the sake of simplicity, the tense of the verb is omitted
the semantic structure in the (D).

25

It is determined whether the final semantic structure
generated as described above is relevant to the standardized
format for semantic structure. By presenting the final
conceptual graph to the user and confirming an intension of
the user, relevance is determined. The (E) presents an
exemplary of the semantic structure generation failure type

data 210 and failure type data format in case that the final semantic structure is not relevant to the standardized format for semantic structure or the semantic ambiguities happen. The semantic structure confirmed by the user is relevant to the standardized format for semantic structure, the conceptual graph is partitioned depending on relation' 211. The (F) presents an example of the partitioned graph for which the conceptual graph in the (D) is partitioned depending on the relation nodes.

After partitioning the conceptual graph, the conceptual graph is indexed when the input sentence from the user is for information supply and the conceptual graph is searched when the input sentence for the user is the query for information request. For the information supply, conceptual graph indexing is generation of the table of relational database including records, each having conceptual pair for the relation and Web document identifiers. The format and contents of the table generated by indexing the partitioned graph in the (F) is shown in Figs. 5A to 5F. When the partitioned graph is for the query representing the user's request information, conceptual graph search is performed so as to retrieve the document semantically nearest to the user's request information represented by the query.

Fig. 6 is a flow chart for an embodiment of user request information and document retrieval procedure in accordance with the present invention.

As shown in Fig. 6, after step 901 of initializing (d-

>the level of the highest node=0, $N \leftarrow$ depth of the partitioned graph), the level of the highest node is compared with the depth of the partitioned graph at step 602 and, if it is less than the depth of the partitioned graph, search for the
5 relation node that belongs to the level (d) of the conceptual graph at step 603.

Subsequently, language characteristic search priority nodes d1, c2 are determined at step 604, it is searched the record of the table related to the relation node (n) depending
10 on the priority rule 608 of the language L1 with respect to the c1 and c2 at step 605, the semantic relevance of each record is computed at step 606, the level of the highest node is increased at step 607, and then the step 602 of comparing the level of the highest node with the depth of the
15 partitioned graph is repeated.

That is, after the initializing step ($d \leftarrow$ the level of the highest node=0, $N \leftarrow$ the depth of the partitioned graph), if there are nodes to be searched at the step 602, it is searched the relation node that belongs to the current graph
20 level. The graph level means location of hierarchical structure of the conceptual graph.

As the conceptual graph in the (D) and the partitioned graph in the (F), the "ATTR" table in Fig. 5A is the relation node of the highest level and the "AGENT" table in Fig. 5B is
25 the relation node of the second highest level. The depth of the conceptual graph is total number of the relation nodes to its terminal node and the depth of the conceptual graph in the

(D) is 4.

After determining the relation node to be searched, the language characteristic semantic relevance priority is determined for a pair of conceptual nodes connected to the relation node. When the relation is embellishing word + embellished word, the conceptual node corresponding to the embellished word has priority. In other cases in the language characteristic semantic relevance priority, the semantic priority is determined by a semantic priority rule uniquely existing for the language. After determining priority, the semantic relevance between each conceptual node and each record of the table for the relation node n is computed. The semantic relevance $S(x,y)$ between concept x and concept y is computed as follows.

$$S(x,y) = \frac{1}{1+d(x,y)} \quad \text{Eq.(1)}$$

where $d(x,y)$ is distance between the nodes in the thesaurus system.

$d(x,y)$, i.e., the distance from the node x to the node y in the thesaurus system, is 0 if the y is one of lower nodes and is computed as the number of edges between the nodes if otherwise. For example, the value of $d(\text{'method'}, \text{'procedure'})$ is 0 when 'procedure' is lower concept than 'method' in the conceptual network. Therefore, the semantic relevance $S(I,Q)$ that a Web document (I) has for a particular query is computed as a sum of the semantic relevances for records indexed for I in the whole tale for the conceptual graph of all sentence

that I has, which can be expressed as follows.

$$S(I,Q) = \sum_l^L \sum_m^M \sum_n^N \frac{1}{d} \{ \delta \cdot s(C_1^l, r_1^{mn}) + (1-\delta) \cdot s(C_2^l, r_2^{mn}) \}$$

After conceptual graph retrieval, the retrieved graphs are aligned in order the values of S(I,Q) and transferred to the response generator 142. The response generator 142 received indexing result, retrieved result or analysis failure type data and generates response sentence and result output format to output them to the user through the information supply interface 16 or information demand interface 12.

As described above, the method of the present invention is implemented as a program that can be stored at a computer readable recording medium, e.g., CD ROM, ROM, RAM, floppy disk, hard disk, optical magnetic disk and etc.

As described above, in the present invention, by describing information to be supplied or requested relevant to the standardized formats for sentence structure and semantic structure for information transaction using the natural language sentence such as English and Korean, efficient retrieval can be performed for tremendous information in the Internet and, by guiding to be relevant to the standardized formats of sentence structure and semantic structure from information generation and description stage, accurate and clear information in the Internet can be generated so that information transaction and transport can be encouraged.

And, in the present invention, by the sentence structure analysis with stratum, the semantic structure analysis in-

depths and semantic relevance computation of the conceptual graph of the semantic structure for the sentence representing the query for information request or supplied information from the user, information of the sentence having different
5 sentence structure but same semantic structure can be retrieved so that demand for transaction of information request or supplied information from the user can be satisfied.

While the present invention has been shown and described with respect to the particular embodiments, it will be
10 apparent to those skilled in the art that many changes and modifications may be made without departing from the spirit and scope of the invention as defined in the appended claims.